(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 099 996 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
16.05.2001 Bulletin 2001/20

(51) Int. Cl.[7]: **G06F 1/00**

(21) Application number: **00309309.3**

(22) Date of filing: **23.10.2000**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **03.11.1999 US 433806**

(71) Applicant:
Ford Global Technologies, Inc.
Dearborn, Michigan 48126 (US)

(72) Inventor: **Pettovello, Primo Mark**
Canton, Michigan 48187 (US)

(74) Representative:
Messulam, Alec Moses
A. Messulam & Co. Ltd.,
43-45 High Road
Bushey Heath, Bushey, Herts WD23 1EE (GB)

(54) Privacy data escrow system and method

(57) The privacy data escrow system (10) includes at least one data provider (12) having a plurality of privacy data records of a plurality of persons. Each privacy data record is associated with a unique person identifier of a person, and each of the at least one data provider (12) having a unique data provider identifier associated therewith. An escrow agent (16) is in communication with the at least one data provider (12) and is operable to receive and store, from the at least one data provider (12), the plurality of person identifiers, and a plurality of unique scrambled person identifiers and data provider identifiers associated with each person identifier (14). A database (20) is in communication with the at least one data provider (12) and is operable to receive and store, from the at least one data provider (12), the plurality of privacy data records, the plurality of scrambled person identifiers associated with the privacy data records, and the data provider identifiers (13). The database (20) is further operable to receive and store, from the escrow agent (16), a unique universal anonymous identifier to replace each scrambled person identifier (18) whereby each privacy data record stored in database is identifiable by a universal anonymous identifier.

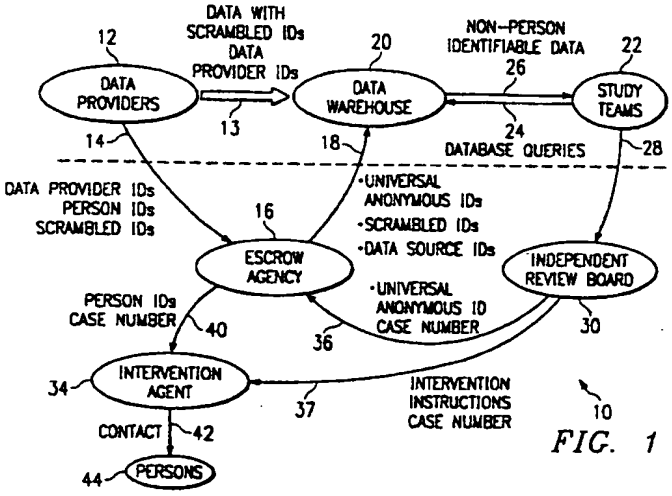

FIG. 1

EP 1 099 996 A1

Description

[0001] This invention relates to computers and computer databases, and more particularly, to a privacy data escrow system and method.

[0002] In today's computer age, nearly every human action leads to the generation, collection and storage of some data. For example, a shopper's grocery or merchandise purchasing habits are collected at the checkout line and stored in databases for future marketing or customer relation purposes. In some instances, sensitive personal data collection leads to privacy issues. For example, financial data are collected whenever a customer applies for credit or a loan, and medical records are maintained for patients for insurance claim purposes. In the latter example, special concerns exist for employees whose employers maintain health care records of its employees. The challenge for employers is to maintain the confidentiality and privacy of employee health medical and claims data, while permitting access to the data for research, analysis, and, in some cases, targeted patient intervention.

[0003] It has been recognised that it is desirable to provide a privacy data escrow system and method to maintain the confidentiality of sensitive personal data such as patient medical records.

[0004] In one aspect of the invention, a privacy data escrow system includes at least one data provider having a plurality of privacy data records of a plurality of persons. Each privacy data record is associated with a unique person identifier of a person, and each of the at least one data provider having a unique data provider identifier associated therewith. An escrow agent is in communication with the at least one data provider and is operable to receive and store, from the at least one data provider, the plurality of person identifiers, and a plurality of unique scrambled person identifiers and data provider identifiers associated with each person identifier. A database is in communication with the at least one data provider and is operable to receive and store, from the at least one data provider, the plurality of privacy data records, the plurality of scrambled person identifiers associated with the privacy data records, and the data provider identifiers. The database is further operable to receive and store, from the escrow agent, a unique universal anonymous identifier to replace each scrambled person identifier whereby each privacy data record stored in database is identifiable by a universal anonymous identifier.

[0005] In another aspect of the invention, a privacy data escrow system includes at least one data provider having a plurality of privacy data records of a plurality of persons, each privacy data record being associated with a unique person identifier of a person, each of the at least one data provider having a unique data provider identifier associated therewith, the at least one data provider being operable to scramble the person identifiers and generate unique scrambled person identifiers therefrom. An escrow agent is in communication with the at least one data provider and is operable to receive and store, from the at least one data provider, the plurality of person identifiers, the associated scrambled person identifiers, and the associated data provider identifier, the escrow agent being operable to generate a unique universal anonymous identifier for each scrambled person identifier. A database in communications with the at least one data provider and is operable to receive and store, from the at least one data provider, the plurality of privacy data records, the plurality of scrambled person identifiers associated with the privacy data records, and the data provider identifier. The database is further operable to receive and store, from the escrow agent, a unique universal anonymous identifier to replace each scrambled person identifier whereby each privacy data record stored in database is identifiable by a universal anonymous identifier.

[0006] In yet another aspect of the invention, a method of maintaining the confidentiality of privacy data includes the steps of associating a unique person identifier with each privacy data record, scrambling the unique person identifier and generating a scrambled person identifier, transmitting the privacy data record and the scrambled person identifier to a database for storage, and transmitting the person identifier with its associated scrambled person identifier to an escrow agency for confidential safekeeping. The escrow agency then generates a universal anonymous identifier for each person identifier and scrambled person identifier, and transmits the universal anonymous identifier and its associated scrambled person identifier to the database.

[0007] The present invention will now be described further, by way of example, with reference to the accompanying drawings, in which:

FIGURE 1 is a simplified data flow diagram of an embodiment of the privacy data escrow system and method according to the teachings of the present invention; and

FIGURE 2 is a more detailed numerical data flow example of an embodiment of the process of separating and scrambling person identifiers from the data according to the teachings of the present invention.

[0008] FIGURE 1 is a simplified data flow diagram of an embodiment of the privacy data escrow system and method 10 according to the teachings of the present invention. Privacy data escrow system 10 obtains sensitive or confidential data from one or more sources or data providers 12. Data providers 12 may be persons, entities, organisations, or companies that has possession of the sensitive data. A data provider 12 may or may not have collected the data itself. In the patient medical records example described above, data providers 12 may be employee health insurance carri-

ers, and the medical records typically include a person identifier such as the social security number of the patient. The medical records may further contain other person identifiable attributes such as name, work and home addresses, work and home phone numbers, and like information. The sensitive data may be insurance claims, clinical records, pharmacy records, occupational health information, worker's compensation information, financial information, personnel information and other data. However, sensitive and confidential data of another nature may be protected by system 10 in the same manner.

[0009] Prior to releasing the sensitive data, data provider 12 separates the person identifier from the rest of the data and scrambles the person identifier. Any data scrambling, encoding or encryption algorithm may be used. The scrambling algorithm may even be a random number generator which uses the person identifier as the seed number. Data provider 12 then transmits or sends a data feed of the data with the scrambled person identifier and a data provider identifier (13) to a database, data management system, or data warehouse 20 for data storage. The sensitive data stored in database 20 is therefore associated only with a data provider identifier and a scrambled person identifier. Data provider 12 also transmits the scrambled person identifier and the associated person identifier along with the data provider identifier (14) to a trusted escrow agent 16 for safe keeping. Other person identifiable attributes which may be used to identify the person are also transmitted to escrow agent 16. Escrow agent 16 therefore possesses a mapping of the scrambled person identifier to the person identifier and other person identifiable attributes. The mapping information may be represented in the form of a table.

[0010] Escrow agent 16 then generates a unique universal anonymous identifier for each person identifier. This universal anonymous identifier is transmitted along with the associated scrambled person identifier and data provider identifier to a database 20. Database 20 thus has sufficient information to map or otherwise associate the scrambled person identifiers to the corresponding universal anonymous identifiers, but not to the person identifiers. In fact, database 20 does not possess any data on the person identifiers or any other data attributes that can be used to identify the person. The universal anonymous identifier is used to reference all data related to a specific person regardless of the identity of the data source or data provider 12. Therefore, each person may be referenced by a unique universal anonymous identifier in database 20 without compromising the confidentiality of the data.

[0011] FIGURE 2 is a more detailed numerical data flow example of an embodiment of the process of separating and scrambling person identifiers from the data according to the teachings of the present invention. The data shown are merely for demonstration purposes and do not resemble actual data.

[0012] Data provider 12 may be a health insurance carrier which has an identifier of "BC11BS," for example. Each data provider 12 in system 10 are uniquely identifiable by a data provider identifier. Data provider 12 has a set of original claims data related to a person identified by person identifier "31313," for example. Typically, the original claims data also includes other person identifiable attributes, such as name, address, phone number, etc. The original claims data in possession by data provider 12 may also include a diagnosis code ("2003"), a procedure code ("J123"), other claims data ("klmnop"), and other assorted data ("qrstuv"). Data provider 12 then applies a scrambling algorithm to methodically alter the person identifier, so that it is now "907432". The scrambled person identifier is unique to each person identifier or person. The scrambled person identifier is transmitted or fed to database 20 with the data provider identifier and the remaining data (for example, diagnosis code, procedure code, other claims data, and other data). Only the person identifier is scrambled or altered from the original. The data is typically transmitted electronically to database 20 via a data feed or EDI (electronic data interchange).

[0013] Data provider 12 also transmits to escrow agency 16 the same scrambled person identifier ("907432"), the data provider identifier ("BC11BS"), the original unscrambled person identifier ("313131"), and all other person identifiable attributes. Using this information, escrow agency 16 either creates a new universal anonymous identifier ("39863211") if the person is new in the system, or looks up the universal anonymous identifier previously assigned to the person. Escrow agent 16 then transmits the mapping from the universal anonymous identifier to the scrambled person identifier to database 20. The data provider identifier may also be sent to database 20. The universal anonymous identifier is substituted for the scrambled person identifier in database 20. Once substituted, all data belonging to a person stored in database 20 are identified by or associated with the same unique universal anonymous identifier.

[0014] It may be seen that the linking or mapping from the universal anonymous identifier to the person identifier provides the key to unlock the anonymity of the data stored in database 20. This key relationship is held in confidence by trusted escrow agency 16 and is kept separate from the data itself stored in database 20. Without the key relationship, the data in database 20 cannot be linked to any person.

[0015] Returning to FIGURE 1, the data in database 20 may be accessed by study teams 22 for analysis, research or other purposes. Database inquiries 24 to database produces non-person identifiable data 26 accessed by study teams 22. If for some reason, members of the study team believes that intervention is required or desirable, then a proposal 28 is made to an independent review board 30. The proposal provides a list of one or more universal anonymous identifiers, a

suggested intervention method or intervention instructions, and evidence supporting the need for the intervention. Independent review board 30 evaluates proposal 28 and makes a decision whether the proposed intervention should be made. If intervention is deemed appropriate, independent review board 30 assigns a unique case number to the intervention and transmits the case number with the intervention instructions to a certified intervention agent 34. Independent review board 30 also sends the case number with the universal anonymous identified(s) to escrow agency 16. Escrow agency 16, upon receipt of the case number and universal anonymous identifier(s), sends the person identifier(s) and other person identifiable attributes associated with the received universal anonymous identifier(s) to intervention agent 34. Intervention agent 34 now has the person identifier(s) and person identifiable attributes along with the intervention instructions. Intervention agent 34 is then able to contact (42) persons 44 as instructed. The intervention method may be telephone calls, written correspondence, physician contact, or any other suitable means. Intervention agent 34 may be an automated process that receives and executes intervention instruction commands from independent review board to automatically prepare a letter or some form of communication for contacting the person. It may be seen that in this procedure, intervention agent 34 does not have access to any data other than person identifiers and intervention instructions, and study teams 22 do not have access to any data other than the anonymous data records.

[0016] It is preferable, in order to achieve and maintain security and integrity of system 10, that all entities operate independently from one another. For example, if the system deals with employee medical insurance claim data, independent review board 30, escrow agency 16, and intervention agent 34 are preferably not related entities of the employer and are able to function independently therefrom. Further, escrow agent 16 is required to safeguard the mapping tables between the universal anonymous identifiers and person identifiers and not release this information to persons or entities without the proper credentials. Additionally, whenever data is electronically transferred via a network, it is preferable that data encryption techniques be used to ensure confidentiality of the data. The data are typically housed in databases such as relational databases, object-oriented databases, relational object-oriented databases and the like.

[0017] It is contemplated by the teachings of the present invention to apply system 10 to any data of a sensitive confidential nature, such as medical health records, medical claim records, pharmacy records, clinical records, lab test results, occupational health information, worker's compensation information, personnel information, genetic information, and personal financial data.

[0018] Although several embodiments of the present invention and its advantages have been described in detail, it should be understood that mutations, changes, substitutions, transformations, modifications, variations, and alterations can be made therein without departing from the teachings of the present invention.

Claims

1. A privacy data escrow system, comprising:

   at least one data provider (12) having a plurality of privacy data records of a plurality of persons, each privacy data record being associated with a unique person identifier of a person, each of the at least one data provider having a unique data provider identifier (13) associated therewith;
   an escrow agent (16) in communication with the at least one data provider (12) and operable to receive and store, from the at least one data provider (12), the plurality of person identifiers and a plurality of unique scrambled person identifiers each having a one-to-one relationship with one of the plurality of person identifiers, and data provider identifiers associated with each person identifier; and
   a database (20) in communication with the at least one data provider (12) and operable to receive and store, from the at least one data provider (12), the plurality of privacy data records, the plurality of scrambled person identifiers associated with the privacy data records, and the data provider identifiers, the database (20) further operable to receive and store, from the escrow agent (16), a unique universal anonymous identifier to replace each scrambled person identifier whereby each privacy data record stored in database is identifiable by a universal anonymous identifier.

2. A system as claimed in claim 1, wherein the escrow agent comprises a mapping table associating the plurality of person identifiers with the universal anonymous identifiers.

3. A system as claimed in claim 1, wherein the privacy data record comprises medical insurance claim data.

4. A system as claimed in claim 1, wherein the privacy data record comprises occupational health data.

5. A system as claimed in claim 1, wherein the privacy data record comprises worker's compensation data.

6. A system as claimed in claim 1, wherein the privacy

data record comprises electronic medical record, clinical data, pharmacy data and medical data.

7. A system as claimed in claim 1, wherein the at least one data provider comprises a data scrambler operable to scramble the person identifier and generate the scrambled person identifier.

8. A system as claimed in claim 1, further comprising an intervention agent in communication with the escrow agent and operable to receive a person identifier therefrom and performing intervention with the person identified by the person identifier.

9. A system as claimed in claim 1, further comprising:

a study team operable to access the privacy data record stored in the database and generating proposed interventions;
an independent review board operable to review and authorise the proposed interventions, the independent review board assigning a case number to an authorised intervention;
the escrow agent receiving the intervention case number and at least one universal anonymous identifier associated with the authorised intervention; and
an intervention agent operable to receive the intervention case number and authorised intervention from the independent review board and further receive, from the escrow agent, at least one person identifier associated with the intervention case number.

10. A method of maintaining the confidentiality of privacy data, comprising:

associating a unique person identifier with each privacy data record;
scrambling the unique person identifier and generating a scrambled person identifier;
transmitting the privacy data record and the scrambled person identifier to a database for storage;
transmitting the person identifier with its associated scrambled person identifier to an escrow agency for confidential safekeeping;
generating, by the escrow agency, a universal anonymous identifier for each person identifier and scrambled person identifier, and transmitting the universal anonymous identifier and its associated scrambled person identifier to the database.

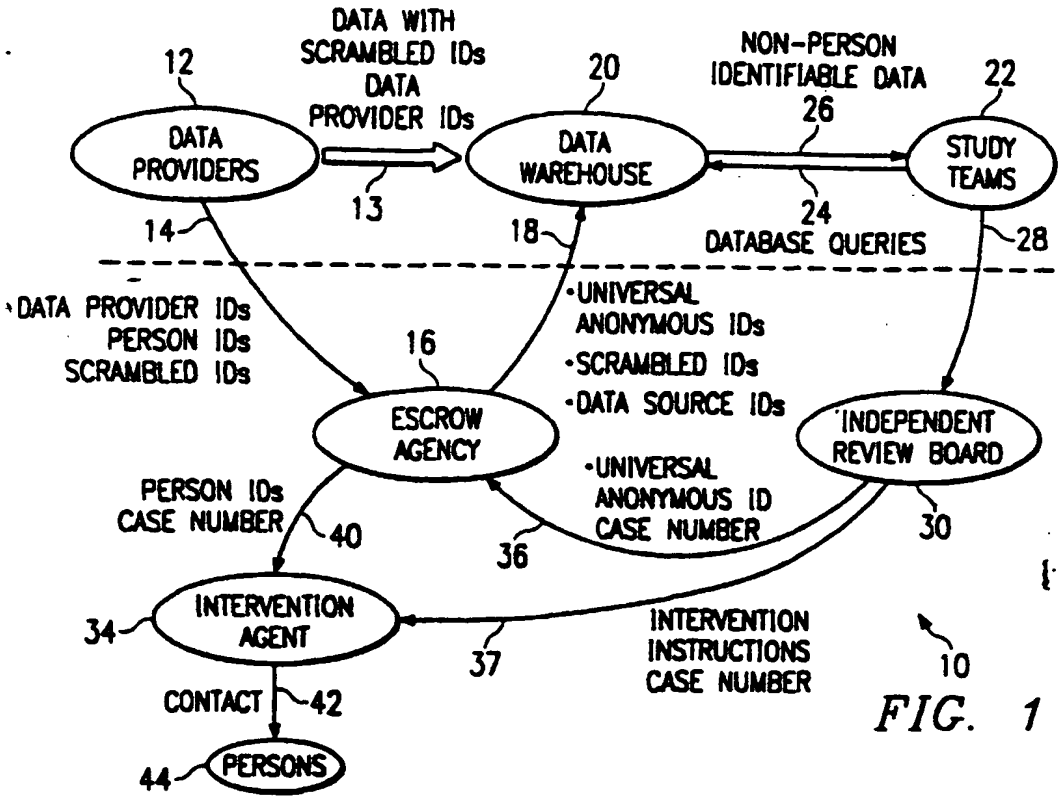
DATA WITH SCRAMBLED IDs
DATA PROVIDER IDs
NON-PERSON IDENTIFIABLE DATA
12
20
22
26
DATA PROVIDERS
DATA WAREHOUSE
STUDY TEAMS
13
24
DATABASE QUERIES
14
18
28
•DATA PROVIDER IDs
PERSON IDs
SCRAMBLED IDs
•UNIVERSAL ANONYMOUS IDs
•SCRAMBLED IDs
•DATA SOURCE IDs
16
ESCROW AGENCY
INDEPENDENT REVIEW BOARD
•UNIVERSAL ANONYMOUS ID CASE NUMBER
PERSON IDs CASE NUMBER
40
36
30
INTERVENTION AGENT
34
37
INTERVENTION INSTRUCTIONS CASE NUMBER
10
CONTACT
42
PERSONS
44

FIG. 1

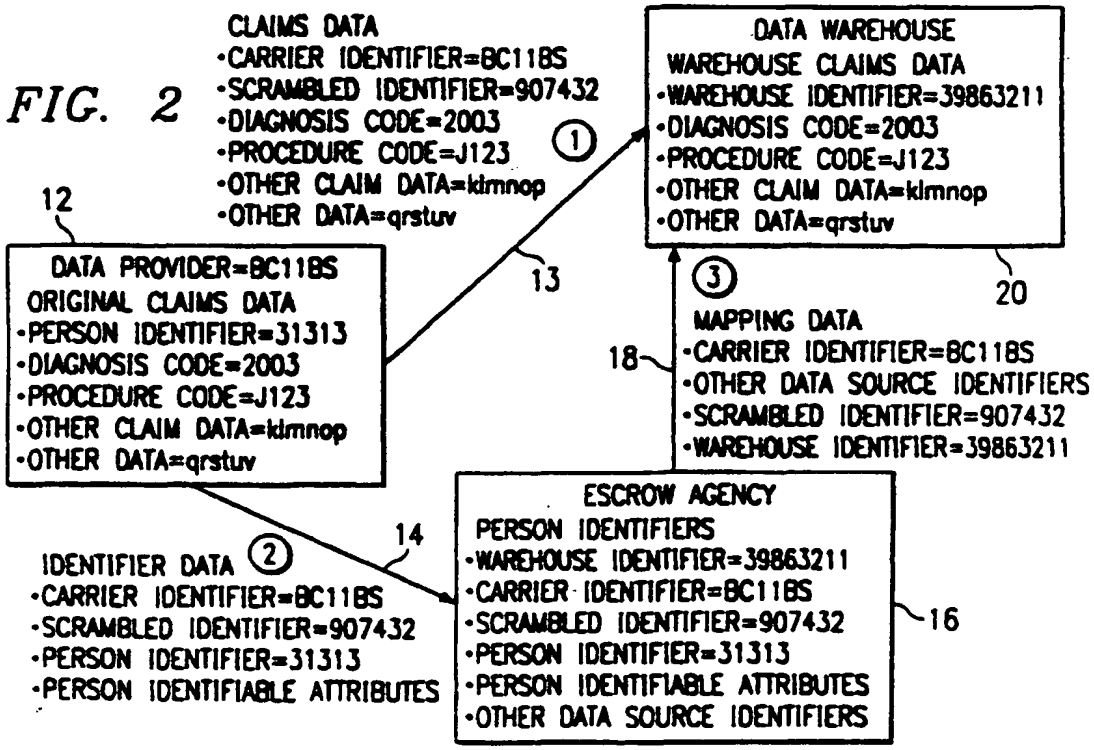
CLAIMS DATA
•CARRIER IDENTIFIER=BC11BS
•SCRAMBLED IDENTIFIER=907432
•DIAGNOSIS CODE=2003
•PROCEDURE CODE=J123
•OTHER CLAIM DATA=klmnop
•OTHER DATA=qrstuv
1
DATA WAREHOUSE
WAREHOUSE CLAIMS DATA
•WAREHOUSE IDENTIFIER=39863211
•DIAGNOSIS CODE=2003
•PROCEDURE CODE=J123
•OTHER CLAIM DATA=klmnop
•OTHER DATA=qrstuv
12
DATA PROVIDER=BC11BS
ORIGINAL CLAIMS DATA
•PERSON IDENTIFIER=31313
•DIAGNOSIS CODE=2003
•PROCEDURE CODE=J123
•OTHER CLAIM DATA=klmnop
•OTHER DATA=qrstuv
13
20
3
MAPPING DATA
18
•CARRIER IDENTIFIER=BC11BS
•OTHER DATA SOURCE IDENTIFIERS
•SCRAMBLED IDENTIFIER=907432
•WAREHOUSE IDENTIFIER=39863211
ESCROW AGENCY
PERSON IDENTIFIERS
•WAREHOUSE IDENTIFIER=39863211
•CARRIER IDENTIFIER=BC11BS
•SCRAMBLED IDENTIFIER=907432
•PERSON IDENTIFIER=31313
•PERSON IDENTIFIABLE ATTRIBUTES
•OTHER DATA SOURCE IDENTIFIERS
16
14
IDENTIFIER DATA
2
•CARRIER IDENTIFIER=BC11BS
•SCRAMBLED IDENTIFIER=907432
•PERSON IDENTIFIER=31313
•PERSON IDENTIFIABLE ATTRIBUTES

FIG. 2

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 00 30 9309

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
A	WO 99 38080 A (HO ANDREW P) 29 July 1999 (1999-07-29) * page 1 - page 13 * * figures 1-5 *	1-10	G06F1/00
A	US 5 956 400 A (VOSKUIL ERIK W ET AL) 21 September 1999 (1999-09-21) * abstract * * column 6, line 15 - column 23, line 24 * * figures 1-4,10 *	1-10	
A	EP 0 884 670 A (INT COMPUTERS LTD) 16 December 1998 (1998-12-16) * abstract * * column 1, line 5 - column 6, line 27 * * figures 1,2 *	1,3-7,10	
A	EP 0 950 972 A (CITICORP DEV CENTER INC) 20 October 1999 (1999-10-20) * column 5, line 8 - column 9, line 14 *	1,10	

TECHNICAL FIELDS SEARCHED (Int.Cl.7)
G06F H04L

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	22 January 2001	Jacobs, P

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

BEST AVAILABLE COPY

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 00 30 9309

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

22-01-2001

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9938080 A	29-07-1999	US 6148342 A	14-11-2000
		AU 2335599 A	09-08-1999
US 5956400 A	21-09-1999	NONE	
EP 0884670 A	16-12-1998	NONE	
EP 0950972 A	20-10-1999	AU 1584499 A	31-05-1999
		AU 1796599 A	31-05-1999
		BR 9806416 A	16-11-1999
		CN 1233804 A	03-11-1999
		EP 0917119 A	19-05-1999
		EP 0917120 A	19-05-1999
		EP 0951158 A	20-10-1999
		EP 0950992 A	20-10-1999
		JP 2000036049 A	02-02-2000
		JP 2000076189 A	14-03-2000
		JP 2000251006 A	14-09-2000
		JP 11250165 A	17-09-1999
		JP 11232348 A	27-08-1999
		WO 9924891 A	20-05-1999
		WO 9924892 A	20-05-1999
		AU 9234698 A	03-06-1999

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82